Complicating to Persuade?

Eduardo Perez-Richet, Delphine Prady

ÉCOLE POLYTECHNIQUE ParisTech

ECOLE POLYTECHNIQUE

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

COMPLICATING TO PERSUADE?

Eduardo PEREZ-RICHET
Delphine PRADY

February 2012

Cahier n° 2012-08

DEPARTEMENT D'ECONOMIE
Route de Saclay
91128 PALAISEAU CEDEX
(33) 1 69333033
http://www.economie.polytechnique.edu/
mailto:chantal.poujouly@polytechnique.edu

Complicating to Persuade? *

Eduardo Perez-Richet† Delphine Prady‡

February 2012

Abstract

This paper addresses a common criticism of certification processes: that they simultaneously generate excessive complexity, insufficient scrutiny and high rates of undue validation. We build a model of persuasion in which low and high types pool on their choice of complexity. A natural criterion based on forward induction selects the high-type optimal pooling equilibrium.When the receiver prefers rejection ex ante, the sender simplifies her report. When the receiver prefers validation ex ante, however, more complexity makes the receiver less selective, and we provide sufficient conditions that lead to complexity inflation in equilibrium.

Keywords: Complexity Inflation, Certification, Persuasion, Strategic Information Transmission, Signaling Games.
JEL classification: D82, C72

> *"Our complex, metastatic, viral systems condemned to the exponential dimension alone (be it that of exponential stability or instability), to eccentricity and indefinite fractal scissiparity, can no longer come to an end. Condemned to an intense metabolism, to an intense internal metastasis, they become exhausted within themselves and no longer have any destination, any end, any otherness, any fatality. They are condemned, precisely, to the epidemic, to the endless excrescences of the fractal and not to the reversibility and perfect resolution of the fateful."*
>
> Jean Baudrillard, The Transparency of Evil, 1993[1].

1 Introduction

Certification systems are often criticized on the ground that they simultaneously generate excessive complexity, insufficient scrutiny by the certifier and high rates of undue validation. For

*Thanks TBA.
†Ecole Polytechnique, mail: eduardo.perez@polytechnique.edu
‡French Treasury, mail: delphine.prady@dgtresor.gouv.fr
[1]We found this less than transparent opening quote in Sokal and Bricmont (1998) which is full (and critical) of such abuses of one's cognitive resources.

example, George Soros commented on the 2008-2010 financial crisis: "The super-boom got out of hand when the new products became so complicated that the authorities could no longer calculate the risks and started relying on the risk management methods of the banks themselves. Similarly, the rating agencies relied on the information provided by the originators of synthetic products. It was a shocking abdication of responsibility."[2] In the case of the Enron scandal, there was the same mix of complex accounting practices and complacency of the auditors. A similar point is also often made about the peer review process in the academic world. The Sokal affair for example can be interpreted along these lines[3].

In this paper we examine this criticism with a theoretical perspective. In order to do that, we build a simple and general model that generates excessive complexity, insufficient scrutiny and undue validations. The model provides clear conditions under which this outcome can be obtained, thus giving us a better understanding of the phenomenon. The central insight is that complexity inflation can only (but does not necessarily) arise when the expected cost of validation for the certifier is low compared to the expected cost of rejection. We also provide sufficient conditions for and examples with complexity inflation.

More generally, our results are useful to understand the use of complexity in any context of persuasion. If a speaker seeks to get a proposition approved by a listener, she can describe her proposition with more or less complex language. It is a widely held belief that in this type of situations artificial complexity may help the speaker. Experts (doctors, mechanics), for example, are sometimes accused of using jargon excessively as a means of persuasion[4]. Our model offers a possible explanation of this phenomenon.

In the model, the product to be certified can be either good or bad. Its properties are

[2]Soros, George (January 22, 2008). "The Worst Market Crisis in 60 Years." *Financial Times (London UK)*.

[3]The main intent behind Sokal's hoax was to denounce the inappropriate use of science by some authors in the humanities. One explanation suggested by Sokal is that it is a way to intimidate the reader with the use of unexplained and excessively complex concepts. This fits our understanding of complexity as barriers to understanding, but the mechanism that leads to insufficient scrutiny in our paper does not involve any irrationality on the receiver's side.

[4]At Bretton Woods, for example, John Maynard Keynes berated American lawyers for writing in Cherokee. He also wrote that "too often lawyers are men who turn poetry into prose and prose into jargon," (Collected Writings, XXVI, p.102).

described in a report whose complexity is chosen by the sender. We assume that for each type of product (good or bad) there is a natural level of complexity from which it is costly to depart. For example, when modeling a new idea in an academic paper, some conceptual elements need to be in place that make it hard to simplify the exposition much. It is also likely that the same idea can be expressed within a more complex framework, but additional layers of complexity are costly to build, be it only because one needs to maintain coherence. The *natural complexity level* of the high type is the one that would be chosen to describe a good product in a world with no bad products and no certification process. We say that there is *complexity inflation* if both types use complexity levels above that natural level. If they both use complexity levels below that level, we talk about *simplification.* Two strengths of our approach are, first that we make very few assumptions on the cost of complexity for the senders of different types, and second that complexity inflation and simplification are relative concepts. We realize that different types of ideas may need to be expressed with a different degree of complexity and this is what the natural complexity level captures. Complication and simplification are departures from this level.

From the point of view of the receiver, higher complexity increases the marginal cost of understanding. It is this property that defines what we mean by complexity: increasing complexity is to make precise understanding more difficult for the receiver. In the case of the financial crisis, the process of securitization made financial products less traceable so that it would be more costly for the rating agencies to retrieve precise information when assessing the associated risk. In the case of academic papers, one could think of complexity as the degree of mathematical sophistication, or of reliance on unfamiliar concepts.

We also assume that complexity is perfectly observable by the receiver. This simplifying assumption is a good approximation as long as the receiver is able to quantify how difficult it is for her to understand the report. After receiving the report, she observes its complexity, chooses her level of understanding, and takes a validation decision. The higher her level of understanding, the easier it is to distinguish good and bad products from one another.

We work under the assumption that it is easy for the low type to imitate the high type, so that separating equilibria are ruled out[5]. Otherwise the low type and the high type would use different levels of complexity and the certifier would identify them and always make the right decision, which is an unrealistic outcome[6]. Among pooling equilibria, we select those that are preferred by the high type. The intuitive reason for doing so is that the high type has a natural leadership in this game since the low type is bound to imitate her in order to stand a chance of being certified. This intuition can be made rigorous through the use of a forward induction based refinement that selects precisely the high-type optimal equilibria.

Given this choice of equilibria, one could be surprised to observe complexity inflation in equilibrium. Indeed, in order to be more easily identified by the certifier, high types should try to facilitate understanding and prefer equilibria where reports are simple. However there is another force that plays in different directions depending on the incentives of the certifier. The certifier can choose how selective to be in her evaluation. It turns out that when she is biased towards validation (inappropriate rejection is more costly than inappropriate validation), less understanding is optimally associated with lower validation cutoffs. Hence the sender might be tempted to increase complexity just to make the certifier less selective, even though it leads to lower levels of understanding. We provide a condition under which the second force overwhelms the first one. This condition takes the form of an upper bound on achievable levels of understanding. When the receiver is biased towards rejection, on the opposite, simplification is doubly beneficial for the high type as it leads to more scrutiny and lower validation cutoffs.

Welfare analysis is slightly delicate in this framework. It is both unclear what our predicted outcome should be compared to, and what should be weighted in the welfare function. To address the second point we evaluate welfare according to a social welfare function that may

[5]In an extension, we provide conditions under which separating equilibria exist but do not satisfy a reasonable refinement.

[6]This point also answers a critique sometimes made about signaling models: what if there were other signals available such as burning money? Wouldn't they be used by the senders to separate? If one is trying to model a situation where mistakes appear to be made in the real world, it must be the case that the senders are pooling on all available signals that they are aware of. Then if the cost of using the different signals available is separable, there is no harm in studying only one of these signals.

put different weights than the receiver on the two types of errors (inappropriate validation and inappropriate rejection), and possibly takes into account a monotonic cost of complexity. To address the first point, we take cheap imitability as a given and therefore compare our predicted outcome to other pooling levels of complexity. With these general welfare functions, we prove that complexity is bad whenever either the receiver has a validation bias and society is more inclined towards rejection, or the receiver has a rejection bias and society is more inclined towards validation.

Related Literature. This paper is naturally connected to the rich literature on strategic information transmission that originated with the seminal papers of Crawford and Sobel (1982) for soft information, and Grossman and Hart (1980); Grossman (1981); Milgrom (1981) for hard information[7]. In our model, information is hard, but the receiver has to pay a learning cost to assimilate its content, while the sender has some control over that cost. A similar idea is present in Bar-Isaac, Caruana and Cuñat (2010) or in the literature on obfuscation (Ellison and Ellison, 2009; Ellison and Wolitzky, 2009; Carlin and Manso, 2010). Caillaud and Tirole (2007) and Perez-Richet (2012) have the same feature but information is equivocal for the sender in the sense that ex ante she is no better than the receiver at predicting the impact of her private information on the decision of the receiver. In all these papers, controlling the learning cost of the receiver is costless for the sender. In our paper, control is exerted through the choice of complexity, the cost of which can vary with minimal restrictions. In this respect, our paper is closer to Dewatripont and Tirole (2005) in which the quality of communication depends on the efforts of both parties. They consider simultaneous and sequential efforts, allowing for different degrees of observability of the sender's effort. However the only case for which they consider sequential efforts is the one where the sender does not know the payoffs of the receiver (which is the same as equivocal information). This precludes the receiver from inferring anything based on the observation of the sender's effort. In contrast, we only look at the sequential process with perfect observability, but in a model where the sender knows the payoff of the

[7] More recently, Kartik (2009) builds a bridge between these two branches of the literature.

receiver, therefore enabling the signaling channel. Another important difference is that in our model it may be costly both to facilitate and to hinder assimilation depending on the choice of complexity. Furthermore, whereas Dewatripont and Tirole (2005) focus on Pareto optimal equilibria, we select equilibria according to a positive criterion based on the idea of forward induction.

Another related paper is Kamenica and Gentzkow (2010), where the sender can commit to a signaling technology before being informed. This feature makes the choice of the sender uninformative by itself, as in Caillaud and Tirole (2007), Perez-Richet (2012). They provide conditions under which the sender may benefit from persuasion because the signaling technology induces the receiver to make a choice that is on average more favorable to the receiver than the one she would have made based on the prior. By comparison, we put more restrictions on the ways a sender can influence the signaling technology faced by the receiver, but the senders acts after being informed, so that her actions may in principle convey some information. In both papers, the actions of the sender more uninformative signaling technologies. In Blume and Board (2010) some types of the sender also choose to be deliberately vague in a very different framework.

This paper is also related to the signaling literature (see Sobel, 2007, for a survey that includes transmission of soft or hard information), first because soft and hard information transmission can be seen as particular cases of signaling, but mostly because the choice of complexity by the sender acts as a signal in our model since it is observable and costly. In particular, we focus on pooling equilibria, and use an equilibrium selection method for the analysis. The most usual selection criteria (the intuitive criterion, D1, divinity) do not help in our model (see Banks and Sobel, 1987; Cho, 1987; Cho and Kreps, 1987; Cho and Sobel, 1990). Instead, we use a selection method described in Umbhauer (1994) which is close to the concept of undefeated equilibria of Mailath, Okuno-Fujiwara and Postlewaite (1993). There are other recent papers that mix costly signaling and an exogenous signaling technology such as Daley and Green (2010) which shows in particular that when the exogenous signaling technology is independent from the costly signal and sufficiently precise, equilibria that are separating in the costly signal no longer survive common refinements.

Finally, our paper contributes to the literature on certification, which is too large for us to summarize here. Skreta and Veldkamp (2009) is, to the best of our knowledge, the only paper in this literature to address the issue of complexity from a theoretical point of view. They present an equilibrium model of the market for ratings of financial products. Complexity is exogenous, and it is defined as the variance in the prior of the rating agency for a given product. They show that an increase in overall complexity can generate rating inflation. In our model, complexity is endogenous, and there is no competition between certifiers.

2 The Model

Sender. A sender S seeks validation from a receiver R. She has a product of type $t \in T \equiv \{L, H\}$. The product is of the high type H with commonly known probability p. Knowing the quality of her product, the sender chooses the *complexity* $\kappa \geq 0$ of the message she sends to the receiver to present her product. If the product is validated by the receiver, she receives a payoff $V_t > 0$. Senders also have the option of not sending any message and receiving 0. We denote this choice by $\kappa = o$. Let $\mathcal{K} = \mathbb{R}_+$ denote the set of possible complexity levels. The action set of the sender is then $\mathcal{K} \cup \{o\}$.

Complexity. For any type of product, there is a continuously differentiable cost of complexity function $C(., t) \geq 0$ defined on $\mathcal{K}$ such that $\lim_{\kappa \to \infty} C(\kappa, t) = \infty$. The cost of opting out is 0. We assume that there exists a unique global minimizer κ_H of $C(\kappa, H)$. It is the *natural complexity* for products of type H. We say that equilibria in which both types use complexity levels above κ_H exhibit *complexity inflation.* Equilibria in which both types use complexity levels below κ_H are *simplifying equilibria.* With this terminology, pooling equilibria in pure strategies are either simplifying or exhibit complexity inflation.

We assume that any complexity level that would be acceptable for the high type if her product were to be validated with probability one would also be acceptable for the low type under the same conditions. That is, letting $\mathcal{K}_t \equiv \{\kappa \in \mathcal{K} : C(\kappa, t) \leq V_t\}$,

$$\mathcal{K}_H \subseteq \mathcal{K}_L. \tag{GI}$$

We call this assumption *Global Imitability* since it implies that the low type can always imitate the high type. $\mathcal{K}_H$ is a compact set that we assume to be nonempty, so that there is some problem to study.

Finally, we say that $C(., H)$ satisfies the *Costly Simplification* assumption if lowering complexity is always costly to the left of the natural level, that is if for every $\kappa \in \mathcal{K}$

$$\kappa \leq \kappa_H \;\Rightarrow\; C_\kappa(\kappa, H) \leq 0. \tag{CS}$$

The latter is not a maintained assumption but it will be useful to provide sufficient conditions for complexity inflation.

Receiver. The receiver's objective is to validate high quality products and reject others. Without loss of generality, we normalize the payoff of the receiver to 0 when she makes the right decision. Undue validation entails a loss $\omega_v > 0$ while undue rejection entails a loss $\omega_r > 0$. We say that the receiver has a *validation bias* if ex ante the expected cost of validation is lower than the expected cost of rejection $(1-p)\omega_v < p\omega_r$. If the inequality is reversed, we say that the receiver has a *rejection bias*. Here bias is not to be understood as a behavioral problem, it is merely a way for us to describe the ex ante preferences of the receiver.

The receiver makes two choices. After receiving the message, she observes its complexity and decides the level of understanding ℓ she wants to reach. The cost of understanding is $c(\ell, \kappa)$, where $c_\ell, c_\kappa, c_{\ell\kappa} > 0$, so that higher complexity raises the marginal cost of understanding[8]. The

[8] The level of understanding ℓ should not be interpreted as an effort. Our assumptions imply that a receiver will choose lower levels of understanding when faced with more complex reports. It does not necessarily mean that she is exerting less effort. Rather it is the cost of understanding that could be interpreted as an effort if one wants to think in these terms.

assumption that total cost increases with complexity is only needed in the section on welfare. We also assume that there exists some $\overline{\ell}$ such that $\lim_{\ell \uparrow \overline{\ell}} c(\ell, \kappa) = \infty$. Let $\mathcal{L} = \left[0, \overline{\ell}\right]$ denote the set of feasible levels of understanding.The receiver then receives a signal σ about the product and decides whether to validate. The informativeness of the signal σ depends on her level of understanding ℓ.

Information. The signal σ is distributed on an interval $\Sigma \subseteq \mathbb{R}$ with conditional density $f(\sigma, \ell, t)$. $f(.)$ is strictly positive on the interior of Σ and continuously differentiable. The associated cumulative density function is denoted by $F(.)$. Let $\lambda(\sigma, \ell) \equiv \frac{f(\sigma,\ell,H)}{f(\sigma,\ell,L)}$ denote the likelihood of the high type for given ℓ and σ.

Definition 1 (Information System)**.** *The pair of density functions $\left\langle f(.,.,L), f(.,.,H)\right\rangle$ on $\Sigma \times \mathcal{L}$ forms an information system if it satisfies the following properties*

(i) $\lambda(\sigma, \ell)$ *is strictly increasing in* σ*.*

(ii) There exists a unique σ_0 *such that for every* ℓ, $\lambda(\sigma_0, \ell) = 1$*.*

(iii) $F(.,.,H)$ *is increasing in* ℓ *in the first-order stochastic dominance order.*

(iv) $F(.,.,L)$ *is decreasing in* ℓ *in the first-order stochastic dominance order.*

(v) $\lambda_\ell(\sigma, \ell) \cdot (\sigma - \sigma_0) \geq 0$ *for every* σ *and every* ℓ*.*

Assumption (i) is the usual monotone likelihood ratio assumption and says that higher signals are more likely if the product is of the high type[9]. Assumption (ii) ensures that the set of signals for which a product is more likely to be good than bad is independent of the level of understanding. Assumption (iii) says that when the product is of the high type, higher scrutiny by the receiver leads to higher signals, and (iv) makes the symmetric statement for low type products. Finally, (v) says that the relative likelihood of signals above σ_0 increases with ℓ while it decreases for signals below σ_0. If (ii) is satisfied, a sufficient condition for *(iii)*,

[9]The assumption could be weakened to weakly increasing without altering our results at the cost of unnecessary technical discussions.

(iv) and (v) is that $f(.,.,H)$ be log-supermodular in (σ, ℓ) and $f(.,.,L)$ be log-supermodular in $(\sigma, -\ell)$. Figure 1 illustrates the properties of information systems and shows the effects of understanding on the log-likelihood. All these assumptions except (ii) simply mean that information is meaningful about types and that understanding works the way we expect it to. Assumption (ii) allows to clarify the discussion but is not crucial. It could be the case that the log-likelihood functions depicted on Figure 1 all rotate around the same point σ_0, but that their value at this point is different from 0, and then our results would hold up to a redefinition of validation and rejection biases. It could also be that the crossing points between two log-likelihood curves differ for different pairs (ℓ, ℓ') and then our results would hold only for strong biases. We show how to extend our results to more general information structures in Appendix C.

For any continuously differentiable symmetric, single-peaked and log-concave density function $\phi(.)$ with support on $\mathbb{R}$, and any $\alpha > 0$, the pair of functions defined by

$$\begin{cases} f(\sigma, \ell, H) = \phi(\sigma - (\alpha + \ell)) \\ f(\sigma, \ell, L) = \phi(\sigma + \alpha + \ell) \end{cases}$$

forms an information system (for a precise definition of single-peakedness and a proof of this claim, see Appendix C). Let $\mathcal{T}$ denote the class of such information systems. An information system in class $\mathcal{T}$ is therefore defined by a pair $\langle \phi(.), \alpha \rangle$. It is such that the density functions corresponding to different levels of understanding are obtained from $\phi(.)$ by translation to the right for the high type, and to the left for the low type. The conditions on $\phi(.)$ are satisfied by any symmetric normal distribution, for example.

3 Equilibrium Definition

We use the concept of perfect Bayesian equilibrium which is equivalent to sequential equilibrium in 2-period games (Fudenberg and Tirole, 1991). Let $\beta(\kappa)$ denote the probability that the

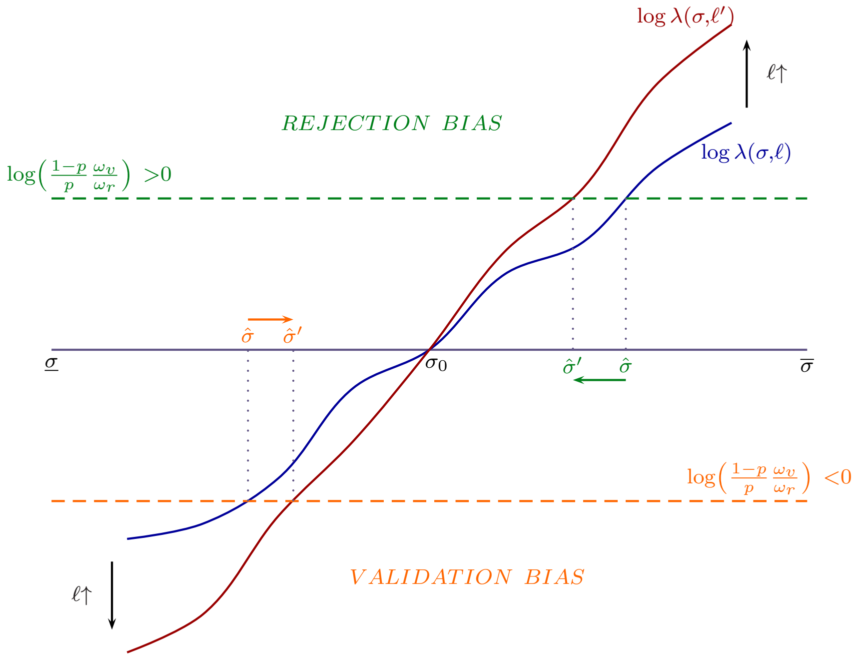

Figure 1: **Understanding Levels and Validation Cutoffs**

receiver assigns to the high type upon observing a report of complexity κ. An equilibrium specifies a strategy of the sender $\kappa(t)$, a belief of the receiver $\beta(\kappa)$, and a strategy of the receiver $\big(\ell(\kappa), v(\ell,\sigma)\big)$, where $v(.)$ is her binary validation decision and takes value 1 in case of validation, such that the receiver makes optimal decisions given her beliefs, the sender reacts optimally to these beliefs and the beliefs are deduced from Bayes rule whenever possible.

Validation. The receiver knows σ, κ and ℓ and takes a validation decision $v \in \{0,1\}$ according to her belief $\beta(\kappa)$ to maximize

$$-v\,\Pr\big(L|\kappa,\sigma,\ell,\beta(.)\big)\omega_v - (1-v)\,\Pr\big(H|\kappa,\sigma,\ell,\beta(.)\big)\omega_r.$$

The relative likelihood of high types given κ, σ and $\beta(.)$ is

$$\Lambda(\kappa,\sigma,\ell,\beta(.)) \equiv \frac{\Pr\big(H|\kappa,\sigma,\ell,\beta(.)\big)}{\Pr\big(L|\kappa,\sigma,\ell,\beta(.)\big)} = \frac{\beta(\kappa)}{1-\beta(\kappa)}\lambda(\ell,\sigma),$$

and the optimal decision of the receiver is to validate when $\Lambda(\kappa,\sigma,\ell,\beta(.)) > \frac{\omega_v}{\omega_r}$, and to reject when the inequality is reversed[10]. Hence a product is validated if the relative likelihood that it is good is higher than the relative cost of validation. The equality condition defines a cutoff $\hat{\sigma}\big(\ell, \beta(\kappa)\big)$ above which products are optimally validated and below which they are optimally rejected. That is, the function $\hat{\sigma}(.,.)$ is implicitly defined as the unique solution to the equation

$$\frac{\beta(\kappa)}{1-\beta(\kappa)}\lambda(\ell,\sigma) = \frac{\omega_v}{\omega_r}. \tag{Valid}$$

Then the following lemma results from a direct application of the implicit function theorem.

Lemma 1.

(i) $\hat{\sigma}\big(\ell, \beta\big)$ *is continuously differentiable in* ℓ *and* β *with derivatives* $\hat{\sigma}_\ell(\ell,\beta) = -\lambda_\ell(\ell,\beta)/\lambda_\sigma(\ell,\beta)$ *and* $\hat{\sigma}_\beta(\ell,\beta) = -\lambda_\sigma(\ell,\beta)/\beta^2$.

(ii) It is increasing in ℓ *if* $\frac{\beta(\kappa)}{1-\beta(\kappa)} > \frac{\omega_v}{\omega_r}$, *and decreasing in* ℓ *if* $\frac{\beta(\kappa)}{1-\beta(\kappa)} < \frac{\omega_v}{\omega_r}$.

(iii) It is decreasing in β.

The comparative statics is illustrated in Figure 1 where we used the log-likelihood and where the belief β is equal to the prior p. When this is true, the two cases in (ii) correspond to whether the receiver has a validation or a rejection bias.

Scrutiny. Given this validation strategy, the problem of the receiver at the information acquisition stage is

$$\max_{\ell\in\mathcal{L}} \; -\beta \Pr\big(\sigma < \hat{\sigma}(\ell,\beta) \;\big|\; \ell, H\big)\omega_r - (1-\beta)\Pr\big(\sigma > \hat{\sigma}(\ell,\beta) \;\big|\; \ell, L\big)\omega_v - c(\ell,\kappa), \tag{Scrut}$$

where we omitted the dependencies of β in κ. This objective function is continuous in ℓ over the compact interval $\mathcal{L}$ implying that there exists an optimal choice of scrutiny level $\ell(\kappa)$ (it is not necessarily unique). The following result is a direct consequence of the monotone selection theorem (Milgrom and Shannon, 1994).

[10] What the receiver does in case of equality is irrelevant for her payoff and has no effect on the behavior of the sender because it is a measure 0 event.

Lemma 2. *Let $\kappa' > \kappa$ be such that $\beta(\kappa) = \beta(\kappa')$. Then for any solution ℓ of the scrutiny problem at κ and any solution ℓ' of the scrutiny problem at κ', we have $\ell' \leq \ell$.*

Proof. When holding $\beta(\kappa)$ constant, the cross derivative with respect to ℓ and κ of the objective function in (Scrut) is $-c_{\ell\kappa} < 0$, and hence it is strictly submodular. The conclusion follows by the monotone selection theorem. □

Complexity. A sender of type t chooses complexity to solve the problem

$$\max_{\kappa \in \mathcal{K}_t \cup \{o\}} \left\{ V_t \Pr\Big(\sigma > \hat{\sigma}\big(\ell^*(\kappa), \beta(\kappa)\big) \mid \ell^*(\kappa), t\Big) - C(\kappa, t) \right\} \mathbb{1}_{\kappa \neq o}. \tag{Cplx}$$

Equilibrium. A perfect Bayesian equilibrium in pure strategies is then defined by a strategy of the sender $\kappa^e(t)$, a belief of the receiver $\beta^e : \mathcal{K} \cup \{o\} \to [0, 1]$, an information acquisition strategy of the receiver $\ell^e : \mathcal{K} \to \mathcal{L}$, and a validation strategy of the receiver $\sigma^e : \mathcal{K} \times \mathcal{L} \to \Sigma$ such that

(i) $\sigma^e(\kappa, \ell)$ solves (Valid).

(ii) $\ell^e(\kappa)$ solves (Scrut).

(iii) $\kappa^e(t)$ solves (Cplx).

(iv) $\beta^e(.)$ is deduced from Bayes rule whenever possible.

4 Analysis

Separating Equilibria. In a separating equilibrium, the receiver infers the quality of an idea from the complexity of the message she receives. Suppose that the high type uses a complexity level κ^0. It must satisfy $C(\kappa^0, H) \leq V_H$. Since the low type has no chance of being validated she can only be choosing not to send any message in equilibrium. For the equilibrium to be separating, the low type must have no incentive to imitate the hight type, that is $V_L - C(\kappa^0, L) \leq 0$. But this is ruled out by the global imitability assumption.

Pooling Equilibria. There always exists a pooling equilibrium in which both types opt out. This equilibrium can be supported with the belief $\beta(\kappa) = p \cdot \mathbb{1}_{\kappa=o}$. In order to characterize other pooling equilibria, we introduce some notations.

Let $\hat{\sigma}(\ell) \equiv \hat{\sigma}(\ell, p)$ be the optimal validation cutoff associated to a level of understanding ℓ under the prior belief. It is the unique solution of the equation in σ

$$\frac{p}{1-p}\lambda(\ell, \sigma) = \frac{\omega_v}{\omega_r}. \tag{ValidPool}$$

Let $\ell^*(\kappa)$ be the set of optimal understanding levels for a given complexity level κ under the prior belief.

$$\ell^*(\kappa) \equiv \arg\max_{\ell \in \mathcal{L}} \; -p \Pr\left(\sigma < \hat{\sigma}(\ell) \middle| \ell, H\right) \omega_r - (1-p) \Pr\left(\sigma > \hat{\sigma}(\ell) \middle| \ell, L\right) \omega_v - c(\ell, \kappa) \tag{ScrutPool}$$

Proposition 1. *All pure strategy equilibria are pooling equilibria, and there always exists a pooling equilibrium in which both types opt out. A triple $e = (\kappa, \ell, \sigma)$ with $\kappa \neq o$ forms a pooling equilibrium if and only if*

(i) $C(\kappa, t) \leq V_t\big(1 - F(\sigma, \ell, t)\big)$, $\forall t \in \{L, H\}$

(ii) $\ell \in \ell^*(\kappa)$.

(iii) $\sigma = \hat{\sigma}(\ell)$

Proof. We have already proved the first claim and the second one is obvious. Let $\tilde{e} = (\tilde{\kappa}, \tilde{\ell}, \tilde{\sigma})$ be a triple satisfying (i), (ii) and (iii). To prove that e is an equilibrium we only need to find a belief that supports it. By definition, this belief must satisfy $\beta(\tilde{\kappa}) = p$. The belief that attributes probability 0 to the high type after observing any $\kappa \neq \tilde{\kappa}$ works. It is clear that any pooling equilibrium must satisfy (ii) and (iii) by definition. It must also satisfy (i) for otherwise at least one of the two types would prefer to opt out. $\square$

Furthermore we have the following corollary of Lemma 1 and Lemma 2.

Lemma 3. *For every selection $\ell(\kappa) \in \ell^*(\kappa)$, we have*

(i) $\ell(\kappa)$ is decreasing in κ.

(ii) $\sigma(\kappa) = \hat{\sigma}(\ell(\kappa))$ is decreasing in κ if $\frac{p}{1-p} > \frac{\omega_v}{\omega_r}$ (validation bias) and increasing in κ if $\frac{p}{1-p} < \frac{\omega_v}{\omega_r}$ (rejection bias).

Let $\mathcal{P}(\kappa)$ be the set of pooling equilibria with complexity level κ. Let $\mathcal{K}^*$ be the set of complexity levels $\kappa \neq o$ such that there exists a pooling equilibrium at κ. The set of equilibria is then $\mathcal{E} = \bigcup_{\kappa \in \mathcal{K}^* \cup \{o\}} \mathcal{P}(\kappa)$. $\mathcal{K}^*$ may be empty, and then there is no active certification process. When $\mathcal{K}^*$ is not empty, all equilibria may be constrained to be below or above κ_H, leading to simplification of complexity inflation in equilibrium. However, this is not the type of effects that we are interested in. Hence in what follows we will use the assumption that κ_H is in the interior of $\mathcal{K}^*$

$$\kappa_H \in Int(\mathcal{K}^*). \tag{A}$$

This assumption is satisfied whenever the stakes are sufficiently high for both types of senders, which is a reasonable assumption for any noteworthy certification system.

Lemma 4. *Keeping everything else fixed, there exist $\underline{V}_L$ and $\underline{V}_H$ such that (A) is satisfied whenever $V_L \geq \underline{V}_L$ and $V_H \geq \underline{V}_H$.*

Proof. If the receiver has a validation bias, the probability that the high type is validated in a pooling equilibrium is bounded below by $\rho_H = 1 - F\big(\hat{\sigma}(\bar{\ell}), 0, H\big)$, and for the low type by $\rho_L = 1 - F\big(\hat{\sigma}(\bar{\ell}), \bar{\ell}, L\big)$. If the receiver has a rejection bias, we can set $\rho_H = 1 - F\big(\hat{\sigma}(0), 0, H\big)$ and $\rho_L = 1 - F\big(\hat{\sigma}(0), \bar{\ell}, L\big)$. Then the bounds of the lemma are given by $\underline{\mathrm{V}}_t = \frac{C(\kappa_H, t)}{\rho_t}$. □

When $\mathcal{P}(\kappa) \neq \emptyset$, it may not be a singleton, and the source of the multiplicity is that there may be several levels of understanding that are optimal for the receiver when both types pool on κ. We can, however, show the following (see Appendix B for a proof).

Lemma 5. *For almost every $\kappa \in \mathcal{K}^*$, $\mathcal{P}(\kappa)$ is a singleton.*

High-Type optimal Equilibria. We choose to focus on the equilibria that are optimal for the high type. Our choice is based on the intuition that since the low type is bound to imitate the high type so as to remain unrecognized, the high type should have some leadership in the choice of equilibrium. This intuition is well captured by a forward induction based equilibrium refinement that we present in Section 6. In what follows, we define these equilibria that we call high-type optimal equilibria (HTE). We denote by $u_S(e, t)$ the payoff of a sender of type t in the equilibrium e.

Definition 2 (High-Type Optimal Equilibrium)**.** *An equilibrium $e \in \mathcal{P}(\kappa)$ is a HTE if for every $\kappa' \neq \kappa$ and every $e' \in \mathcal{P}(\kappa')$,*

$$u_S(e, H) \geq u_S(e', H)$$

There always exists a HTE. The existence proof is slightly technical and is given in Appendix A. Its main steps are to show first that $\mathcal{K}^*$ is a compact set, and second that the function that to a complexity level $\kappa \in \mathcal{K}^*$ associates the highest payoff that a high type sender can obtain in an equilibrium $e \in \mathcal{P}(\kappa)$ is well defined and upper-semicontinuous, and therefore attains its maximum on $\mathcal{K}^*$. Clearly, if κ_0 is its maximizer, then there exists an equilibrium $e \in \mathcal{P}(\kappa_0)$ which is a HTE.

Proposition 2. *There exists a HTE, and if $\mathcal{K}^* \neq \emptyset$ there exists a HTE which is not the opt-out equilibrium.*

Proof. See Appendix A. □

If the receiver has a rejection bias, increasing κ has two effects that both go against the high type: (i) it makes the receiver sloppier, which is bad because the high type is less likely to be identified, and (ii) it makes the receiver more selective which is bad for both types because there is always a chance that the receiver will receive a bad signal. Therefore we should not expect to have complexity inflation in a HTE. If the receiver has a validation bias, however, the two effects play in opposite directions. When complexity is increased, the receiver still becomes sloppier, but she also becomes less selective. This observation is formalized in the following proposition. The last point of the proposition is made precise and proved in the next section which provides sufficient conditions for complexity inflation.

Proposition 3. *Assume that* (A) *holds. Then, if the receiver has a rejection bias, any HTE is a simplifying equilibrium. If the receiver has a validation bias, however, a HTE may exhibit complexity inflation.*

Proof. Here we prove only the first point of the proposition. Let $e' = (\kappa', \ell', \sigma')$ be an equilibrium. Suppose $\kappa' > \kappa_H$. By (A), there exists another pooling equilibrium $e_H = (\kappa_H, \ell_H, \sigma_H)$. The payoff of the high type in an equilibrium e is given by $u_S(e, H) = V(1 - F(\sigma, \ell, H)) - C(\kappa, H)$. By Lemma 3, we have $\ell_H \geq \ell'$ and $\sigma_H \leq \sigma'$. Because $F(\sigma, \ell, H)$ is increasing in σ and decreasing in ℓ and $C(\kappa_H, H) < C(\kappa', H)$ we have $u_S(e', H) < u_S(e_H, H)$, hence e' cannot be a HTE. □

5 Sufficient Conditions for Complexity Inflation

In the general case, we can write a sufficient condition for complexity inflation. For any given information system, it is possible, although not always easy, to verify whether it is satisfied.

Proposition 4. *Assume that* (A) *holds, that* $C(., H)$ *satisfies* (CS), *and that the receiver has a validation bias. If for every* $\kappa \leq \kappa_H$ *and every pooling equilibrium* $e = (\kappa, \ell, \sigma) \in \mathcal{P}(\kappa)$ *we have*

$$\frac{-\lambda_\ell(\sigma, \ell)}{\lambda_\sigma(\sigma, \ell)} \geq \frac{-F_\ell(\sigma, \ell, H)}{f(\sigma, \ell, H)}, \tag{Infl}$$

then any HTE exhibits complexity inflation.

If $\ell^*(\kappa)$ were a well behaved continuously differentiable functions, we would simply write that the objective function of the high type is given by

$$V_H\Big(1 - F(\hat{\sigma}(\ell^*(\kappa)), \ell^*(\kappa), H)\Big) - C(\kappa, H),$$

and differentiate it with respect to κ to look for a maximum. Then (Infl) would be the condition needed for this derivative to be strictly positive everywhere to the left of κ_H implying that the optimal choice of κ for the high type is to the right of κ_H. This is essentially what the proof does while dealing with the technicalities implied by the lack of smoothness of the problem (see Appendix B for the complete proof).

The inconvenient of this sufficient condition is that it not obvious, just by looking at it, whether it is satisfied by any information system. In order to find examples of information systems that lead to complexity inflation, and thus prove the last point of Proposition 3, we focus on information systems in class $\mathcal{T}$. It is then easy to provide a sufficient condition that can be checked immediately and comes in the form of an upper bound on the ability to discriminate between types at any achievable understanding level. Figure 2 illustrates how to verify whether the condition is satisfied for a given information system $\big\langle \phi, \alpha \big\rangle$.

Proposition 5. *Suppose that the information system is in class $\mathcal{T}$ so that it is defined by a pair $\big\langle \phi(.), \alpha \big\rangle$, that (A) holds, that $C(., H)$ satisfies (CS), and that the receiver has a validation bias. If*

$$\alpha + \bar{\ell} \leq \frac{\Gamma}{2}, \tag{Infl'}$$

where Γ is the unique positive solution of the equation $\phi(\Gamma) = \frac{1-p}{p}\frac{\omega_v}{\omega_r}\phi(0)$, then any HTE exhibits complexity inflation.

The proof in Appendix B shows that for information systems in class $\mathcal{T}$, (Infl') implies (Infl).

Example 1 (The Normal Case)**.** *If $\phi(.)$ is the density of the standard normal distribution, then the likelihood function is given by $\lambda(\ell, \sigma) = \exp(2\sigma(\alpha + \ell))$, the validation cutoff is $\hat{\sigma}(\ell) = \frac{1}{2(\alpha+\ell)} \log\left(\frac{1-p}{p}\frac{\omega_v}{\omega_r}\right)$ and the sufficient condition for complexity inflation under validation bias is*

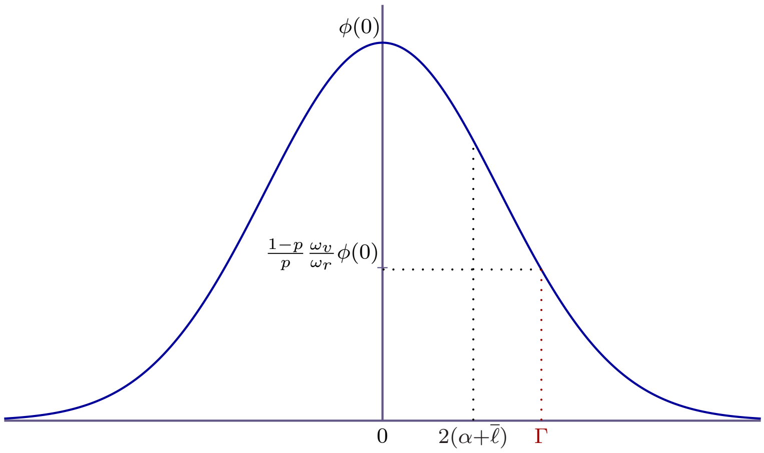

Figure 2: **A Sufficient Condition for Complexity Inflation**

$\alpha + \bar{\ell} < \sqrt{-\frac{1}{2}\log\left(\frac{1-p}{p}\frac{\omega_v}{\omega_r}\right)}.$

6 Equilibrium Refinements

The Usual Criteria. Interestingly, many of the most commonly used refinement concepts are useless for this model, in the sense that all the pure strategy pooling equilibria we identified pass the corresponding test. The source of the problem is that the action space of the receiver in our model is two dimensional since it involves the choice of a scrutiny level and a validation cutoff. Furthermore, the two types of the sender have coinciding interest on one of these dimensions – they would both like the receiver to lower her validation cutoff – and conflicting interests on the other dimension – the good type is better off when the receiver exerts more scrutiny, whereas the bad type prefers sloppy receivers. Criteria such as D1 and divinity rely on identifying for a certain deviation κ' from an equilibrium e the set $\mathcal{A}_t(\kappa', e)$ of actions by the receiver that are preferred to the equilibrium action by each type t, and then put more weight on those types t for whom the set $\mathcal{A}_t(\kappa, e)$ is maximal for the inclusion order. But in our model, the two dimensional action space of the receiver and the nature of conflict imply that in general $\mathcal{A}_H(\kappa', e) \setminus \mathcal{A}_L(\kappa', e) \neq \emptyset$ and $\mathcal{A}_L(\kappa', e) \setminus \mathcal{A}_H(\kappa', e) \neq \emptyset$. The reason is that even though the two types have converging interests in the selectivity dimension (the choice of the cutoff σ), they have conflicting interests in the scrutiny dimension (the choice of ℓ). Hence the nesting condition needed for these notions to operate effectively is generally not satisfied. This is illustrated in Figure 3 in which (a) illustrates the general situation. Point e is the pair of actions chosen by the receiver in the equilibrium under consideration. H and L are the indifference curves of the high type and the low type that go through e. H' is the set of receiver action pairs that compensate the high type for a deviation from the complexity level used in e to another level κ', L' is the same for the low type. In the particular case of figure (a), the change in complexity is beneficial for the high type so that she would accept a lower utility if she could use κ', but it is costly for the low type. $\mathcal{A}_H(\kappa', e)$ is the set of action pairs to the left of H' and $\mathcal{A}_L(\kappa', e)$ is the set of action pairs to the left of L'. In the case of figure (a), none of these sets is a subset of the other and therefore there is no restriction on the equilibrium beliefs that the receiver can hold after observing a deviation to κ'. It is possible that some deviations lead to case (b) or (c). In case (c), $\mathcal{A}_H(\kappa', e) \subset \mathcal{A}_L(\kappa', e)$ and the equilibrium belief that a deviation to κ' is due to the low type with probability 1 is compatible with divinity and D1. The only case in which these refinements can eliminate some equilibria is when a deviation κ' leads to situation (b), which is possible if some deviations are very costly for the low type but not so for the high type. However, the most general case is that of situation (a) in which divinity and D1 are not helpful. In Appendix D, we show that all equilibria can be supported by a belief such that they satisfy the intuitive criterion. Furthermore, if the low type and the high type are sufficiently undistinguishable then all equilibria can be supported by a belief such that they satisfy D1 and divinity.

Umbhauer–Undefeated Equilibria. Instead we use the refinement notion of Umbhauer (1994) which is a variant on the notion of undefeated equilibria of Mailath, Okuno-Fujiwara and Postlewaite (1993) (the two notions were developed independently). It allows us to uniquely select the high-type optimal pooling equilibrium. In order to show that, we briefly redefine this notion for our particular framework and then prove the result. All along we adopt the termi-

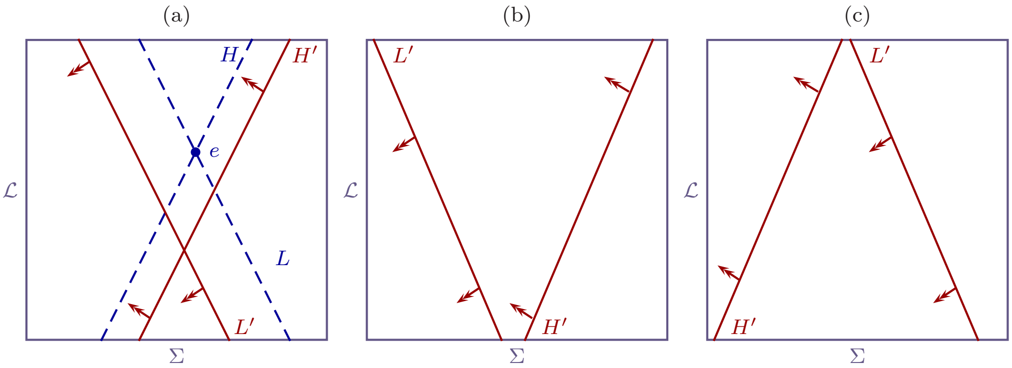

Figure 3: **Senders' preferences over the receiver's actions.** The arrows show the direction of increasing payoff for the two types.

nology of Mailath, Okuno-Fujiwara and Postlewaite (1993) with the definitions of Umbhauer (1994), and we will clarify where the two notions differ.

The refinement works as follows. Consider a perfect Bayesian equilibrium and a complexity level κ_0 that is not used by any type in equilibrium. Suppose there is an alternative perfect Bayesian equilibrium in which some types use κ_0, and that at least one of these types is strictly better off in the alternative equilibrium than in the proposed equilibrium. The test requires the beliefs of the receiver when observing κ_0 in the original equilibrium to be consistent with the set of types that use κ_0 in the alternative equilibrium and are strictly better off. In other words, when the receiver observes the out-of-the-equilibrium-path complexity level κ, she should interpret it as an attempt from the types who benefit more from another equilibrium in which they use κ to switch to this equilibrium. In the following definition $e \equiv \big(\kappa(.), \sigma(.), \ell(.), \beta(.)\big)$ is the equilibrium being tested, and $e' \equiv \big(\kappa'(.), \sigma'(.), \ell'(.), \beta'(.)\big)$ is another equilibrium used to perform the test. For any complexity level κ_0, let

$$T(\kappa_0, e' \succ e) \equiv \{t \in T : \kappa'(t) = \kappa_0 \text{ and } u_S(e', t) > u_S(e, t)\}$$

be the set of types that use κ_0 in e' and strictly prefer the alternative equilibrium e' to the candidate equilibrium e. In this definition, we denote by $\beta(t|\kappa)$ the probability assigned to type t by the receiver after observing κ.

Definition 3 (Umbhauer–Defeated Equilibrium). *Let e and e' be two equilibria of the game. Then e' defeats e if there exists a complexity level κ_0 such that*

(UD1) $\{t \in T : \kappa(t) = \kappa_0\} = \emptyset$ *and* $T(\kappa_0, e' \succ e) \neq \emptyset$.

(UD2) There exists $t \in T(\kappa_0, e' \succ e)$ such that $\beta(t|\kappa_0)$ cannot be derived from a Bayesian update of the prior assigning probability 0 of sending κ_0 to any $t' \notin T(\kappa_0, e' \succ e)$ and uniform positive probability to any type $t' \in T(\kappa_0, e' \succ e)$.

(UD1) says that the level κ_0 is not used in e, but that there are some types who use it in e' and strictly prefer e' to e. (UD2) says that the beliefs in the original equilibrium must be inconsistent with this set of types acting as in e'. If the beliefs are consistent, then the original equilibrium is deemed reasonable and passes the test.

The difference between this modified notion of undefeated equilibrium and that of Mailath, Okuno-Fujiwara and Postlewaite (1993) is that they would require every type using the complexity level κ_0 in the alternative equilibrium to at least weakly prefer the alternative equilibrium. The game we study in this paper is a good illustration of why the notion of Umbhauer–undefeated equilibrium may be preferable. Suppose indeed that the candidate equilibrium e is not a high-type optimal equilibrium and that the alternative equilibrium e' is high-type optimal. Let κ_0 be the complexity level used by both types in e'. Suppose finally that the low type is worse off in e'. The high type may still try to play the out of the equilibrium path complexity level κ_0 in e to try to switch to e', for if this move is correctly interpreted by the receiver as coming from the high type, the low type will have to comply and switch to e' or be identified as a low type and never be approved. In this sense the notion of Umbhauer–undefeated equilibrium perfectly captures the vague intuition that the high type should have some leadership in the choice of equilibrium since the low type has to imitate her in order not to be recognized and punished[11].

[11]Another concept that relates to this intuition, is that of core mechanism in Myerson (1983). We do not pursue this connection here as it would require to recast the model in terms of mechanism design by the sender who would then be an informed principal. We thank Joel Sobel for suggesting this connection.

Proposition 6. *The set of HTE is exactly the set of Umbhauer–undefeated equilibrium.*

Proof. Let e be any pooling equilibrium that is not a high-type optimal equilibrium, and let e' be a high-type optimal pooling equilibrium in which both types use the complexity level κ_0. By definition, H is strictly better off in e'. Suppose first that L is worse off in s. Then the beliefs in the original equilibrium e about the out of the equilibrium path complexity level κ_0 must satisfy $\beta(\kappa_0) = 1$ for that equilibrium not to be defeated by e'. But if that is the case, e cannot be an equilibrium for the high type would profitably deviate to κ_0 and be validated with probability one. If both types are strictly better off in e', then for e not to be defeated we must have $\beta(\kappa) = p$. But then again the high type would be better off by deviating to κ_0 since κ_0 maximizes the payoff of the high type when the belief of the receiver is p. This proves both that the set of Umbhauer-undefeated equilibria is a subset of the set of HTE and the reverse inclusion. □

7 Welfare

We've created a setup with multiple (pooling) equilibria. We've argued that some of these equilibria, the HTE are more natural than others, and we've shown that HTE could exhibit either simplification or complexity inflation, depending essentially but not uniquely on the type of bias of the receiver. One way to do welfare analysis in this context is to ask how the HTE performs compared to other pooling equilibria. In a sense, the question we ask is what is the price of forward induction.

First we look at the payoff of the receiver, which may be a reasonable way to assess the performance of the certification system. We denote by $u_R(e)$ the payoff of the receiver in an equilibrium $e \in \mathcal{P}(\kappa)$. The envelope theorem of Milgrom and Segal (2002) implies that an increase in equilibrium complexity is always detrimental to the receiver.

Proposition 7. *Let $e_0 \in \mathcal{P}(\kappa_0)$ and $e_1 \in \mathcal{P}(\kappa_1)$ be two equilibria such that $\kappa_0 < \kappa_1$. Then $u_R(e_0) > u_R(e_1)$.*

Proof. Let $u_R(\kappa)$ denote the value function of the optimization problem in (ScrutPool). If $e \in \mathcal{P}(\kappa)$, we have $u_R(e) = u_R(\kappa)$. The objective function in (ScrutPool) is continuously differentiable in ℓ and κ and its derivative with respect to κ is equal to $-c_\kappa(\ell, \kappa) < 0$. Then the envelope theorem of Milgrom and Segal (2002) implies that $u_R(\kappa)$ is almost everywhere differentiable on $\mathcal{K}_H$ and wherever the derivative exists $u_R'(\kappa) = -c_\kappa(\ell(\kappa), \kappa)$ for any selection $\ell(\kappa) \in \ell^*(\kappa)$. Furthermore, we can write that for any selection $\ell(\kappa) \in \ell^*(\kappa)$,

$$u_R(\kappa_1) - u_R(\kappa_0) = \int_{\kappa_0}^{\kappa_1} -c_\kappa(\ell(k), k)dk.$$

Hence $u_R(\kappa_0) > u_R(\kappa_1)$. □

For the second result, we assume that society only cares about the two types of errors in validation, with weights $\bar{\omega}_v$ and $\bar{\omega}_r$ that possibly differ from those of the receiver. The prior belief π of society may differ as well from the receiver's prior. Society is more biased towards rejection than the receiver if $\frac{(1-\pi)\bar{\omega}_v}{\pi\bar{\omega}_r} \geq \frac{(1-p)\omega_v}{p\omega_r}$, and more biased towards validation if the opposite inequality holds.

Proposition 8. *If the receiver has a validation bias and society is more biased towards rejection than the receiver, then welfare is decreasing in complexity. If the receiver has a rejection bias and society is more biased towards validation than the receiver, then welfare is decreasing in complexity.*

Proof. The welfare associated with an effort level ℓ and the corresponding optimal cutoff $\hat{\sigma}(\ell)$ is given by

$$W(\ell) \equiv -\pi\bar{\omega}_r F(\hat{\sigma}(\ell), \ell, H) - (1-\pi)\bar{\omega}_v\big(1 - F(\hat{\sigma}(\ell), \ell, L)\big),$$

an its derivative with respect to ℓ is given by

$$-\hat{\sigma}'(\ell)\Big(\pi\bar{\omega}_r f(\hat{\sigma}(\ell), \ell, H) - (1-\pi)\bar{\omega}_v f(\hat{\sigma}(\ell), \ell, L)\Big) + \Big(-\pi\bar{\omega}_r F_\ell(\hat{\sigma}(\ell), \ell, H) + (1-\pi)\bar{\omega}_v F_\ell(\hat{\sigma}(\ell), \ell, L)\Big).$$

The second term is positive by points (iii) and (iv) in Definition 1. For the first term, the definition of $\hat{\sigma}$ implies that

$$p\omega_r f(\hat{\sigma}(\ell), \ell, H) = (1-p)\omega_v f(\hat{\sigma}(\ell), \ell, L),$$

and therefore $\pi\bar{\omega}_r f(\hat{\sigma}(\ell), \ell, H) - (1-\pi)\bar{\omega}_v f(\hat{\sigma}(\ell), \ell, L)$ is negative whenever society is more biased towards rejection than the receiver and positive otherwise. Since $\hat{\sigma}'(.)$ is positive when the receiver has a validation bias and negative in case of a rejection bias, we have proved that $W(\ell)$ is increasing in ℓ in each of the cases mentioned in the proposition. Now if $e_0 \in \mathcal{P}(\kappa_0)$ and $e_1 \in \mathcal{P}(\kappa_1)$ are two equilibria of the game with $\kappa_0 \leq \kappa_1$, we have $\ell_0 \geq \ell_1$ implying that welfare is higher in e_0 than in e_1. □

The most interesting part of this result is that there are cases in which complexity may be good from a social point of view. We cannot conclude for sure because complexity always has the detrimental effect of lowering the level of understanding of the receiver. But it may also compensate for the differences in preferences between society and the receiver. For example, if society and the receiver both have a validation bias, but the bias of society is stronger, more complexity makes the receiver more lenient to the benefit of society. The other case is if both have a rejection bias and society is more strongly biased. In this case, more complexity makes the receiver more selective to the benefit of society.

8 Extensions

Mixed Strategies. We have looked at equilibria in pure strategies. We could, however, allow mixed strategies for the receiver without perturbing the analysis. First, separating equilibria would be discarded in the same way. Then the receiver would never mix at the decision stage, since after having observed complexity, and knowing her understanding level, there is unique optimal cutoff. The only case in which the receiver would want to mix in equilibrium at the scrutiny stage is in the cases of equilibria such that $\mathcal{P}(\kappa)$ is not a singleton (we know that it is almost never the case by Lemma 5). And then the receiver would mix between understanding levels in $\ell^*(\kappa)$, without changing her equilibrium payoff. The effect on the sender's payoff would be to convexify the graph over $\mathcal{K}^*$ of the correspondence that gives the equilibrium payoff of the sender of type $t = L, H$ for any complexity level that can be supported as an equilibrium $u_S^*(\kappa, t)$. All our results would carry through. On the other hand, allowing the sender to play mixed strategies is much more problematic and maybe not that useful, and we do not try to do it.

Information Structure. Most assumptions we make about the information structure are completely natural, or technical but unfalsifiable and do not need to be discussed in detail. Our only important simplifying assumption is (ii) in Definition 1, and it is not crucial for the main results. It is important because it allows us to make clear predictions for any combination of prior beliefs and preferences for the receiver. Under a more general assumption, we would only be able to make predictions for strong validation and rejection biases. These more general assumptions and their consequences are explained in more detail in Appendix C, as well as the role of more technical assumptions such as differentiability of the density functions.

Separating Equilibria. We have worked under the Global Imitability assumption that rules out separating equilibria. Our motivation for doing so is the casual observation that people seem to make mistakes in reality and that, since it would not be the case in any separating equilibrium, these equilibria should be discarded as unrealistic. We think that this argument has some weight, and global imitability does not seem unreasonable by itself in its basic assumption that complexity is cheap for low types. However we also use a selection argument that can help us generalize the global imitability assumption to all the cases where a separating equilibrium may exist but would not be selected by our criterion. If the global imitability assumption is not satisfied, we can still find the high-type optimal pooling equilibria (HTPE). Then let $u_S^{HTPE}(H)$ be the payoff of the high type sender in any HTPE. Then we have the following result which needs no proof.

Proposition 9. *If for every $\kappa \in \mathcal{K}_H \setminus \mathcal{K}_L$, $V_H - C(\kappa, H) < u_S^{HTPE}(H)$, then no separating* *equilibrium is high-type optimal, that is the set of HTE of the game is exactly the set of HTPE.*

Cheap Complexity. If we assume that the cost of complexity is 0 for both types, and limit available complexity levels to $\mathcal{K} = [0, \overline{\kappa}]$, we obtain the following results. First, note that the global imitability assumption is satisfied, and all equilibria are pooling equilibria. Furthermore there is a pooling equilibrium for every complexity level, $\mathcal{K}^* = \mathcal{K}$. When the receiver has a rejection bias, every HTE uses the lowest possible complexity level $\kappa = 0$. Under a validation bias, the complexity level may be positive. Under the same conditions as Proposition 5, every HTE uses the highest possible complexity level $\kappa = \overline{\kappa}$.

Unobservable Complexity. One may prefer to assume that complexity is unobservable, but still affects the choice of a level of understanding by the receiver. The difficulty with this approach is that it implicitly assumes that the receiver is not able to invert her cost function in order to infer complexity from her cost cost of choosing a certain level of understanding. However that may be a plausible description of behavior. In this case we obtain the following results. First, the two types may use different complexity levels. Under a rejection bias, the high type will always simplify with respect to her natural complexity level. The low type will face two opposing forces and she may complicate or simplify. If the conditions in Proposition 5 are satisfied, she will complicate with respect to her natural complexity level κ_L. Under a validation bias, the low type will always complicate with respect to κ_L, whereas the high type will face two opposing forces, but will simplify for sure if the conditions in Proposition 5 are satisfied.

9 Final Remarks

We have proposed a model and a method of equilibrium selection to analyze the choice of complexity in contexts of persuasion. Our results shed light on the conditions that may lead to complexity inflation in equilibrium. On the belief side, optimism of the receiver is likely to lead to complexity inflation. On the preferences side, low costs of undue validation relative to undue rejection is another factor that may lead to complexity inflation. This suggests that collusion between rating agencies and the financial industry and optimism about the quality of financial products, for example, create a favorable environment for excessive complexity. Another factor that may lead to complexity inflation is the reputation of the sender if it translates into stronger priors.

Appendix A Existence of HTE

We start with a useful lemma.

Lemma 6. *$\mathcal{K}^*$ is a compact set.*

Proof. We show that $\mathcal{K}^*$ is closed and bounded. Because any $\kappa \in \mathcal{K}^*$ satisfies (i) in Proposition 1, it must also satisfy $C(\kappa, H) \leq V_H$, hence $\mathcal{K}^* \subseteq \mathcal{K}_H$ is bounded. Then the function $\mu : \mathcal{K} \times \mathcal{L} \to \mathbb{R}$ defined by

$$\mu(\kappa, \ell) \equiv \min\Big\{V_H\left(1 - F(\hat{\sigma}(\ell), \ell, H)\right) - C(\kappa, H)\ ,\ V_L\left(1 - F(\hat{\sigma}(\ell), \ell, L)\right) - C(\kappa, L)\Big\}$$

is continuous as the minimum of two continuous functions. Note also that it is bounded upward by $M = \max\big\{V_L, V_H\big\}$. Now consider the correspondence $\hat{\mu} : \kappa_H \to \mathbb{R}$ defined by $\hat{\mu}(\kappa) \equiv \mu(\kappa, \ell^*(\kappa))$. $\ell^*(.)$ is upper hemicontinuous by the maximum theorem (see for example Aliprantis and Border, 2006, Theorem 17.31). Hence $\hat{\mu}(.)$ is itself upper hemicontinuous. But then the image of a closed set by its lower inverse is itself closed (Aliprantis and Border, 2006, Lemma 17.4), where the lower inverse is defined by

$$\hat{\mu}^{low}(A) \equiv \big\{\kappa \in \mathcal{K} : \hat{\mu}(\kappa) \cap A \neq \emptyset\big\},$$

for any $A \subset \mathbb{R}$. Now Proposition 1 implies that

$$\mathcal{K}^* = \hat{\mu}^{low}\left([0, M]\right),$$

and hence $\mathcal{K}^*$ is closed. $\square$

Proof of Proposition 2 (Existence of HTE). First note that if $\mathcal{K}^* = \emptyset$, the unique equilibrium is the opt out equilibrium and it is clearly a HTE. Suppose from now on that $\mathcal{K}^* \neq \emptyset$. Then any equilibrium $e \in \mathcal{P}(\kappa)$ for $\kappa \in \mathcal{K}^*$ must give a payoff higher than 0 to the high type. If all such equilibria give a payoff of 0, then they are all HTE as well as the opt out equilibrium. Hence from now on we also assume that there exists $\kappa \in \mathcal{K}^*$ and $e \in \mathcal{P}(\kappa)$ such that $u_S(e, H) > 0$. We consider the function $g : \mathcal{K}^* \to \mathbb{R}$ defined by

$$g(\kappa) \equiv \max_{e \in \mathcal{P}(\kappa)} u_S(e, H).$$

The set of equilibria in $\mathcal{P}(\kappa)$ is isomorphic to the set $\ell^*(\kappa) \times \hat{\sigma}(\ell^*(\kappa))$ which is compact by application of the maximum theorem (Aliprantis and Border, 2006, Theorem 17.31). Since the function that to a pair (ℓ, σ) associates the payoff of the high type at a given κ is continuous, it attains its maximum by the extreme value theorem, and therefore $g(.)$ is well defined on $\mathcal{K}^*$. It is also easy to see that $g(.)$ is upper semicontinuous. Then the extension of the extreme value theorem to semicontinuous functions (Aliprantis and Border, 2006, Theorem 2.43) implies that $g(.)$ attains its upper bound on the compact $\mathcal{K}^*$. Let κ_0 be a complexity level that maximizes $g(.)$, and e_0 an equilibrium in $\mathcal{P}(\kappa_0)$ such that $u_S(e_0, H) = g(\kappa_0)$. Then e_0 is a HTE. Indeed, if e is another equilibrium in $\mathcal{P}(\kappa)$ with $\kappa \neq \kappa_0$, then by definition $u_S(e_0, H) \geq u_S(e, H)$. $\square$

Appendix B Additional Proofs

Proof of Lemma 5. Because the objective function in (ScrutPool) is strictly submodular in ℓ and κ, the monotone selection theorem implies that every selection from $\ell^*(\kappa)$ is decreasing in κ. Since the objective function is continuous and the set $\mathcal{L}$ is compact, we can define the two extremal selections $\overline{\ell}^*(\kappa) \equiv \max \ell^*(\kappa)$ and $\underline{\ell}^*(\kappa) \equiv \min \ell^*(\kappa)$. They are both monotonic in κ and therefore almost everywhere continuous. Suppose that there exists an open interval

$I \subset \mathcal{K}_H$ such that for every $\kappa \in I$, $\ell^*(\kappa)$ is not a singleton. Then, for every $\kappa \in I$, $\underline{\ell}^*(\kappa) < \overline{\ell}^*(\kappa)$. Since $\overline{\ell}^*(.)$ and $\underline{\ell}^*(.)$ are almost everywhere continuous, we can choose I so that they are both continuous on I. But then let $\tilde{\kappa} \in I$ and define the selection $\ell(\kappa) = \begin{cases} \underline{\ell}^*(\kappa) \text{ if } \kappa < \tilde{\kappa} \\ \overline{\ell}^*(\kappa) \text{ if } \kappa \geq \tilde{\kappa} \end{cases}$. By continuity of $\underline{\ell}^*(.)$ at $\tilde{\kappa}$, there exists $\kappa < \tilde{\kappa}$ such that $\ell(\kappa) = \underline{\ell}^*(\kappa) < \overline{\ell}^*(\tilde{\kappa}) = \ell(\tilde{\kappa})$. But then $\ell(.)$ is not decreasing in κ which contradicts the monotone selection theorem. Hence $\ell^*(\kappa)$ is almost everywhere a singleton, and since to each ℓ is associated a single σ by the function $\hat{\sigma}(.)$, $\mathcal{P}(\kappa)$ is almost everywhere a singleton. □

Proof of Proposition 4 – Sufficient Condition for Inflation 1. Let $e_0 \in \mathcal{P}(\kappa_0)$ with $\kappa_0 < \kappa_H$, and let $e_H \in \mathcal{P}(\kappa_H)$. We know that $\ell_H \leq \ell_0$ and $\sigma_H \leq \sigma_0$. Let $e(\kappa)$ be a selection of equilibria such that $e(\kappa_0) = e_0$ and $e(\kappa_H) = e_H$. Then Lemma 3 implies that $\ell(\kappa)$ and $\sigma(\kappa)$ are both monotonic. Therefore they must be almost everywhere differentiable, and everywhere left and right differentiable. Let $\psi(\kappa) \equiv u_S(e(\kappa), H)$. The properties of $F(.,.,H)$ and $C(.,H)$ imply that $\psi(.)$ is almost everywhere continuous, almost everywhere differentiable and everywhere left and right differentiable, and therefore we can define the function $\psi'(\kappa)$ which is almost everywhere equal to

$$- V_H \ell'(\kappa)\Big(\hat{\sigma}'(\ell(\kappa))f(\sigma(\kappa), \ell(\kappa), H) + F_\ell(\sigma(\kappa), \ell(\kappa), H)\Big) - C_\kappa(\kappa, H). \tag{1}$$

From (ValidPool), it follows that

$$\hat{\sigma}'(\ell) = \frac{-\lambda_\ell(\hat{\sigma}(\ell), \ell)}{\lambda_\sigma(\hat{\sigma}(\ell), \ell)}.$$

Since $\ell'(\kappa) \leq 0$, (Infl) then implies that the first term in (1) is always positive for $\kappa < \kappa_H$. Because $C(.,H)$ satisfies (CS), we have $C_\kappa(\kappa, H) < 0$ for every $\kappa < \kappa_H$. Hence $\psi'(\kappa) < 0$ for almost every $\kappa \in (\kappa_0, \kappa_H)$. Now remark that every jump in $\psi(.)$ is due to a jump in $\ell(.)$. Suppose for example that there is a jump at $\kappa \leq \kappa_H$, due to a jump of ℓ from ℓ^- to ℓ^+, with

$\ell^- > \ell^+$ since $\ell(.)$ is decreasing. Then the size of the jump is given by

$$\Delta\psi(\kappa) = -V_H \int_{\ell^-}^{\ell^+} \Big(\hat{\sigma}'(\ell) f(\hat{\sigma}(\ell), \ell, H) + F_\ell(\hat{\sigma}(\ell), \ell, H)\Big) d\ell,$$

and this expression is positive by (Infl). Hence all jumps of $\psi(.)$ on $[\kappa_0, \kappa_H]$ are upward. Therefore we can write that

$$\psi(\kappa_H) - \psi(\kappa_0) \geq \int_{\kappa_0}^{\kappa_H} \psi'(\kappa) d\kappa > 0,$$

implying that $\psi(\kappa_H) > \psi(\kappa_0)$. But then e_0 is not a HTE. □

Proof of Proposition 5 – Sufficient Condition for Inflation 2. First note that for every (σ, ℓ), $-F_\ell(\sigma, \ell, H)/f(\sigma, \ell, H) = 1$, and

$$\frac{-\lambda_\ell(\sigma, \ell)}{\lambda_\sigma(\ell, \sigma)} = \frac{\phi'(\sigma - (\alpha + \ell))\phi(\sigma + \alpha + \ell) + \phi(\sigma - (\alpha + \ell))\phi'(\sigma + \alpha + \ell)}{\phi'(\sigma - (\alpha + \ell))\phi(\sigma + \alpha + \ell) - \phi(\sigma - (\alpha + \ell))\phi'(\sigma + \alpha + \ell)},$$

which is greater than 1 if and only if $\phi'(\sigma + \alpha + \ell) \geq 0$. For (Infl) to hold it is therefore sufficient that for every $\sigma \leq \hat{\sigma}(\bar{\ell})$ and every $\ell \leq \bar{\ell}$, $\phi'(\sigma + \alpha + \ell) \geq 0$, which is true as long as $\hat{\sigma}(\bar{\ell}) \leq -(\alpha + \bar{\ell})$. Because $\lambda(\sigma, \bar{\ell})$ is increasing in σ, the latter is true if and only if

$$\lambda\big(\hat{\sigma}(\bar{\ell}), \bar{\ell}\big) = \frac{1-p}{p}\frac{\omega_v}{\omega_r} \leq \lambda\left(-(\alpha + \bar{\ell}), \bar{\ell}\right) = \frac{\phi\left(-2(\alpha + \bar{\ell})\right)}{\phi(0)},$$

and using the symmetry of $\phi(.)$, this is equivalent to (Infl'). □

Appendix C Information Systems

Here we show some properties of information systems that were mentioned in the presentation of the model without proofs, and we show how the main results of the paper extend to strong biases if we allow for a more general information structure than the one assumed in the paper.

We start with some definitions for which we adopt the terminology of Athey (2002).

Definition 4 (Single-Crossing Properties)**.**

1. *A real function $g(y)$ satisfies the single-crossing property SC1 if there exists y_0' and y_0'' such that $g(y) < 0$ for all $y < y_0'$, $g(y) = 0$ for all $y_0' < y < y_0''$, and $g(y) > 0$ for all $y > y_0''$. It satisfies the single-crossing property about a point y_0, denoted SC1(y_0), if $y_0 = y_0' = y_0''$.*

2. *A real function $h(x, y)$ defined on $X \times Y \subset \mathbb{R}^2$satisfies the single-crossing in two variables SC2 in (x, y) if for every $x' > x$, $g(y) = h(x', y) - h(x, y)$ satisfies SC1. SC2(y_0) is defined analogously.*

Going back to the notations of the paper, we can show the following proposition for two density functions $f(\sigma, \ell, H)$ and $f(\sigma, \ell, L)$ on $\Sigma \times \mathcal{L}$. Note that some of the main results in this proposition are adaptations of well known results that we prove here for the sake of clarity and completeness.

Proposition 10. *If $f(\sigma, \ell, H)$ satisfies SC2 in (ℓ, σ) and $f(\sigma, \ell, L)$ satisfies SC2 in $(-\ell, \sigma)$, then the following conditions are satisfied*

(a) $F(., ., H)$ increases with ℓ in the first-order stochastic dominance order.

(b) $F(., ., L)$ decreases with ℓ in the first-order stochastic dominance order.

(c) There exist σ^- and σ^+ in the closure of Σ such that for every $\ell' > \ell$ and every $\sigma > \sigma^+$, $\lambda(\sigma, \ell') > \lambda(\sigma, \ell)$, while for every $\ell' > \ell$ and every $\sigma < \sigma^-$, $\lambda(\sigma, \ell') < \lambda(\sigma, \ell)$.

If in addition there exists σ_0 such that, for every ℓ, $\lambda(\sigma_0, \ell) = 1$, then we can add that $\sigma^- \leq \sigma_0 \leq \sigma^+$.

If instead of SC2 we require that $f(\sigma, \ell, H)$ is strictly log-supermodular in (ℓ, σ), and $f(\sigma, \ell, L)$ is strictly log-supermodular in $(-\ell, \sigma)$, a stronger condition than SC2, and there exists σ_0 such that, for every ℓ, $\lambda(\sigma_0, \ell) = 1$ then we can replace (c) by

(c′) $\lambda(\sigma, \ell)$ satisfies SC2(σ_0).

Finally, if $f(.,.,H)$ and $f(.,.,L)$ are also continuously differentiable, then we can write that $\lambda_\ell(\sigma,\ell)\cdot(\sigma-\sigma_0)\geq 0$ for every σ and ℓ.

Proof. An implication of SC2 for $f(\sigma,\ell,H)$ is that for every $\ell'>\ell$, and every σ it must be true that either for every $\sigma'\geq\sigma$, $f(\sigma,\ell',H)\geq f(\sigma,\ell'L)$, or for every $\sigma\leq\sigma'$, $f(\sigma,\ell',H)\leq f(\sigma,\ell'L)$, where in each case the inequality is strict on a subset of Σ of positive measure. In the first case we can write that

$$1-F(\sigma,\ell',H)=\int_\sigma^{\overline{\sigma}} f(x,\ell',H)dx>\int_\sigma^{\overline{\sigma}} f(x,\ell,H)dx=1-F(\sigma,\ell,H),$$

that is $F(\sigma,\ell',H)\leq F(\sigma,\ell,H)$. And in the second case we can write that

$$F(\sigma,\ell',H)=\int_{\underline{\sigma}}^{\sigma} f(x,\ell',H)dx<\int_{\underline{\sigma}}^{\sigma} f(x,\ell,H)dx=F(\sigma,\ell,H).$$

This proves (a), and (b) is proved similarly. To prove (c), let $\Sigma_H(\ell,\ell')$ denote the closed interval of values of σ such that $f(\sigma,\ell',H)-f(\sigma,\ell,H=0)$, and let $\Sigma_L(\ell,\ell')$ be defined similarly. Finally let $\Sigma_0\equiv\Sigma_H\cup\Sigma_L$ where $\Sigma_t\equiv\bigcup_{\ell'>\ell}\Sigma_t(\ell,\ell')$. Σ_0 is a subset of $Cl(\Sigma)$, the closure of Σ, and therefore we can define $\sigma^+\equiv\sup\Sigma_0$ and $\sigma^-\equiv\inf\Sigma_0$ in $Cl(\Sigma)$. If Σ is unbounded, σ^- and σ^+ may be infinite. Then by SC2, we have for every $\sigma>\sigma_+$ and every $\ell'>\ell$, $f(\sigma,\ell',H)>f(\sigma,\ell,H)$ and $f(\sigma,\ell',L)<f(\sigma,\ell,L)$ and therefore $\lambda(\sigma,\ell')>\lambda(\sigma,\ell)$. A symmetric statement can be made for $\sigma<\sigma^-$.

If in addition there exists σ_0 such that, for every ℓ, $\lambda(\sigma_0,\ell)=1$ it is clear that σ_0 is a crossing point of any two functions $\lambda(\sigma,\ell)$ and $\lambda(\sigma,\ell')$ so that $\sigma_0\in\Sigma_0$ implying that σ_0 is between σ^- and σ^+.

If instead of SC2 we assume strict log-supermodularity, which implies SC2, then it is easy to see that $\lambda(\sigma,\ell)$ is also strictly log-supermodular in (σ,ℓ). Then we can show that for every $\ell'>\ell$, $\lambda(\sigma,\ell')-\lambda(\sigma,\ell)$ crosses 0 at most once. Suppose instead that $\sigma_1<\sigma_2$ are two distinct zeros of this function. Then we have $\lambda(\sigma_1,\ell')-\lambda(\sigma_1,\ell)=\lambda(\sigma_2,\ell')-\lambda(\sigma_2,\ell)$ which contradicts strict log-supermodularity. Since σ_0 is a zero of this function by assumption, it must be the only one. This implies (c'). The last point of the proposition is obvious. □

The property we need for our results to hold with strong biases is (c). If it holds, we can define $\mu^+ \equiv \lambda\big(\sigma^+, \overline{\ell}\big)$, and $\mu^- \equiv \lambda\big(\sigma^-, \overline{\ell}\big)$. Since $\lambda(\sigma, \ell)$ is increasing in σ, we have $\mu^- \leq \mu^+$. We say that the receiver has a *strong rejection bias* if $\frac{(1-p)\omega_v}{p\omega_r} > \mu^+$ and a *strong validation bias* if $\frac{(1-p)\omega_v}{p\omega_r} < \mu^-$. Then all the results of Section 4 hold if we replace the expression validation (respectively, rejection) bias by strong validation (respectively, rejection) bias. Furthermore, if we assume the existence of σ_0 such that for every ℓ, $\lambda(\sigma_0, \ell) = 1$, then $\mu^- \leq 1 \leq \mu^+$, implying that a strong rejection bias implies a rejection bias (and symmetrically for validation).

The proposition shows that (c) is implied by the two single-crossing properties for $f(.,.,H)$ and $f(.,.,L)$, which are natural assumptions about the signaling technology. Point (ii) in Definition 1 allows us to make predictions for any combination of beliefs and preferences of the receiver, but it is not crucial.

Next, we show that signaling technologies in class $\mathcal{T}$ defined in the presentation of the model are indeed information systems.

Definition 5 (Single-Peakedness). *A real function $\phi(.)$ is single-peaked if there exists a unique maximum m such that $w < x < m < y < z$ implies $\phi(w) < \phi(x) < \phi(m)$ and $\phi(m) > \phi(y) > \phi(z)$.*

Proposition 11. *For any continuously differentiable symmetric single-peaked density function $\phi(.)$ with support on $\mathbb{R}$, such that $\phi'(x)/\phi(x)$ is strictly decreasing, and any $\alpha > 0$, the pair of functions defined by*

$$\begin{cases} f(\sigma, \ell, H) = \phi(\sigma - (\alpha + \ell)) \\ f(\sigma, \ell, L) = \phi(\sigma + \alpha + \ell) \end{cases}$$

forms an information system

Proof. Note first that the unique maximum of ϕ is reached at 0 by symmetry. We prove that the functions $f(.,.,H)$ and $f(.,.,L)$ satisfy the five points of Definition 1.

(i) $\lambda_\sigma(\ell,\sigma)$ has the same sign as

$$\phi'(\sigma-(\alpha+\ell))\phi(\sigma+\alpha+\ell)-\phi(\sigma-(\alpha+\ell))\phi'(\sigma+\alpha+\ell).$$

If $-(\alpha+\ell)\le\sigma\le\alpha+\ell$, then $\phi'(\sigma-(\alpha+\ell))\ge 0$ and $\phi'(\sigma+\alpha+\ell)\le 0$ with strict inequalities for almost every σ in $[-(\alpha+\ell),\alpha+\ell]$. Therefore $\lambda_\sigma(\sigma,\ell)>0$ for almost every σ in $[-(\alpha+\ell),\alpha+\ell]$ and $\lambda(\sigma,\ell)$ is strictly increasing in σ on $[-(\alpha+\ell),\alpha+\ell]$. If $\sigma>\alpha+\ell$, then the fact that $\phi'(x)/\phi(x)$ is strictly decreasing on $\mathbb{R}_+$ implies $\lambda_\sigma(\ell,\sigma)\ge 0$. The last case can be dealt with by symmetry.

(ii) Clearly $\sigma_0=0$.

(iii) We have $F(\sigma,\ell,H)=\int_{-\infty}^{\sigma}\phi(x-(\alpha+\ell))dx$, and therefore $F_\ell(\sigma,\ell,H)=-\phi(\sigma-(\alpha+\ell))<0$.

(iv) We have $F(\sigma,\ell,L)=\int_{-\infty}^{\sigma}\phi(x+\alpha+\ell)dx$, and therefore $F_\ell(\sigma,\ell,L)=\phi(\sigma+\alpha+\ell)>0$.

(v) $\lambda_\ell(\ell,\sigma)$ has the same sign as

$$-\phi'(\sigma-(\alpha+\ell))\phi(\sigma+\alpha+\ell)-\phi(\sigma-(\alpha+\ell))\phi'(\sigma+\alpha+\ell).$$

It is positive if $\sigma\ge-(\alpha+\ell)$ because $\phi(.)$ is decreasing on $\mathbb{R}_+$. If $0\le\sigma\le\alpha+\ell$, it is by symmetry equal to

$$\phi'(\alpha+\ell-\sigma)\phi(\sigma+\alpha+\ell)-\phi(\alpha+\ell-\sigma)\phi'(\sigma+\alpha+\ell),$$

which is negative because $\phi'(x)/\phi(x)$ is decreasing on $\mathbb{R}_+$. The results for $\sigma\le 0$ are obtained by symmetry.

$\square$

Appendix D Refinements

This section shows the limitations of the intuitive criterion, D1 and divinity for our game. In order to state our results, we start by introducing two measures of distinguishability between types. For a real subset S, we denote its closure by $Cl(S)$, its interior by $Int(S)$ and its frontier by $\partial S \equiv Cl(S) \setminus Int(S)$.

We define *Incentive Distinguishability* of the two types as

$$\Delta^{inc} \equiv \sup_{\kappa \in \mathcal{K}_H} \left| \frac{C(\kappa, H)}{V_H} - \frac{C(\kappa, L)}{V_L} \right| .$$

The division by V_t amounts to a normalization that makes comparisons possible between types. The supremum is taken over $\mathcal{K}_H$ because any complexity level outside of this set cannot be individually rational for the high type. The supremum exists since $\mathcal{K}_H$ is compact.

Informational Distinguishability is defined as

$$\Delta^{info} \equiv \sup_{\sigma \in \Sigma} \; |F(\sigma, 0, H) - F(\sigma, 0, L)| .$$

It is the distance between the two cumulative distribution functions at the minimal scrutiny level.

Finally, the *Minimal Rejection Wedge* across scrutiny levels, when the selectivity cutoff is chosen optimally, is given by

$$\Omega^{rej} \equiv \inf_{\ell \in \mathcal{L}} \; |F(\hat{\sigma}(\ell), \ell, H) - F(\hat{\sigma}(\ell), \ell, L)| .$$

Because the monotone likelihood ratio property (i) in Definition 1 implies that $F(\sigma, 0, H) < F(\sigma, 0, L)$ for all $\sigma \in Int(\Sigma)$, $\Omega^{rej} > 0$ unless it is optimal to accept or reject all products for some levels of efforts (this is possible only for strong validation or rejection biases).

We say that the setup satisfies the *Sufficient Indistinguishability Condition* if[12]

$$\Delta^{info} + 2\Delta^{inc} < \Omega^{rej}. \tag{SIC}$$

Proposition 12. *For every equilibrium $e \in \mathcal{E}$, there exists a supporting belief $\beta(.)$ for e such that the equilibrium (e, β) satisfies the intuitive criterion. Furthermore, if the Sufficient Indistinguishability condition is satisfied, then there exists a supporting belief such that e satisfies the intuitive criterion, D1 and divinity.*

Proof. First we can eliminate deviations of the sender to a level κ' that is not in $\mathcal{K}_L$ since they are dominated by the choice of opting out for both types (by the Global Imitability assumption). Consider an equilibrium e, and a possible deviation $\kappa' \in \mathcal{K}_L$. Suppose $\kappa' \in \mathcal{K}_L \setminus \mathcal{K}_H$. Then $\mathcal{A}_H(\kappa', e) = \emptyset$, $\mathcal{A}_L(\kappa', e) \neq \emptyset$ and the belief $\beta(\kappa') = 0$ is consistent with divinity. Suppose now that $\kappa' \in \partial\mathcal{K}_H \cap \partial\mathcal{K}_L$, then $\mathcal{A}_t(\kappa', e) = \{(\underline{\sigma}, \ell) : \ell \in \mathcal{L}\}$ for $t = L, H$, where $\underline{\sigma} = \inf \Sigma$ may be $-\infty$ (choosing $\underline{\sigma}$ means that every product is validated), and the belief $\beta(\kappa') = 0$ is consistent with divinity. If instead $\kappa' \in \partial\mathcal{K}_H \cap Int(\mathcal{K}_L)$, then $\mathcal{A}_H(\kappa', e) \subset \mathcal{A}_L(\kappa', e)$ and the belief $\beta(\kappa') = 0$ is consistent with divinity. Suppose finally that $\kappa' \in Int(\mathcal{K}_H)$. Then $\mathcal{A}_H(\kappa', e)$ and $\mathcal{A}_L(\kappa', e)$ are both nonempty, and the set of actions of the receiver that make the deviation equivalent to the initial equilibrium for type t is the set of pairs (σ, ℓ) such that

$$V_t\big(1 - F(\sigma, \ell, t)\big) - C(\kappa', t) = V_t\big(1 - F(\sigma_e, \ell_e, t)\big) - C(\kappa_e, t) \tag{2}$$

Because $F_\ell(\sigma, \ell, H) < 0$ and $F_\ell(\sigma, \ell, L) > 0$, these sets are given by a strictly increasing function $\sigma_H(\ell)$ for H and by a strictly decreasing function $\sigma_L(\ell)$ for L as in Figure 3. Then there are three possible cases depicted in the figure. In case (a), neither $\mathcal{A}_H(\kappa', e)$ or $\mathcal{A}_L(\kappa', e)$ is a subset of the other and the belief $\beta(\kappa') = 0$ is consistent with divinity. In case (c), we have $\mathcal{A}_H(\kappa', e) \subset \mathcal{A}_L(\kappa', e)$ and the belief $\beta(\kappa') = 0$ is again consistent.

Hence the only problematic case is (b), and we will show that it cannot occur under (SIC).

[12]This property can be satisfied only if $\Omega^{rej} > 0$.

For that, note that (b) occurs when $\sigma_L(0) \leq \sigma_H(0)$. Going back to (2), $\sigma_L(0)$ and $\sigma_H(0)$ satisfy the following equation respectively for $t = L, H$

$$F(\sigma_t(0), 0, t) = F(\sigma_e, \ell_e, t) + \frac{C(\kappa_e, t)}{V_t} - \frac{C(\kappa', t)}{V_t}.$$

Hence the difference $F(\sigma_L(0), 0, L) - F(\sigma_H(0), 0, H)$ is equal to

$$F(\sigma_e, \ell_e, L) - F(\sigma_e, \ell_e, H) + \frac{C(\kappa_e, L)}{V_L} - \frac{C(\kappa_e, H)}{V_H} + \frac{C(\kappa', H)}{V_H} - \frac{C(\kappa', L)}{V_L} \geq \Omega^{rej} - 2\Delta^{inc}.$$

Then, by (SIC), we have

$$F(\sigma_L(0), 0, L) - F(\sigma_H(0), 0, H) > \Delta^{info}.$$

By definition, $\Delta^{info} \geq F(\sigma_H(0), 0, L) - F(\sigma_H(0), 0, H)$. Hence we have shown that $F(\sigma_L(0), 0, L) > F(\sigma_H(0), 0, L)$. Since $F(., 0, L)$ is increasing, we have $\sigma_H(0) < \sigma_L(0)$ so that case (b) never occurs, and this concludes the proof for divinity, and for D1 since it is implied by divinity. For the intuitive criterion, we need to show that it can be satisfied even when (SIC) does not hold. This is easy since whenever we are in case (b), $\mathcal{A}_H(\kappa', e)$ and $\mathcal{A}_L(\kappa', e)$ are both non-empty, and in the other cases, the intuitive criterion is implied by divinity. □

References

Aliprantis C., Border K. (2006), "Infinite Dimensional Analysis: A Hitchhiker's Guide," 3rd ed, Springer, Berlin Heidelberg New York.

Athey S. (2002), "Monotone Comparative Statics under Uncertainty," *Quarterly Journal of Economics*, 67, 1, 187-223.

Banks J., Sobel J. (1987), "Equilibrium Selection in Signaling Games," *Econometrica*, 55, 647-661.

Bar-Isaac H., Caruana G., Cuñat V. (2010), "Information Gathering and Marketing," *Journal of Economics and Management Strategy*, 19, 2, 375-401.

Blume A., Board O. (2010), "Intentional Vagueness," working paper.

Caillaud B., Tirole J. (2007), "Consensus Building: How to Persuade a Group," *American Economic Review*, 97(5), 1877-1900.

Carlin B. I., Manso G. (2010), "Obfuscation, Learning and the Evolution of Investor Sophistication," *Review of Financial Studies*, 25(3).

Cho I.-K. (1987), "A Refinement of the Sequential Equilibrium Concept," *Econometrica*, 55, 1367-1389.

Cho I.-K., Kreps D. (1987), "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102, 179-221.

Cho I.-K., Sobel J. (1990), "Strategic Stability and Uniqueness in Signaling Games," *Journal of Economic Theory*, 50, 381-413.

Crawford V., Sobel J. (1982), "Strategic Information Transmission," *Econometrica*, 50-6, 1431-51.

Daley B., Green B. (2010), "Market Signaling with Grades," working paper.

Dewatripont M., Tirole J. (2005), "Modes of Communication," *Journal of Political Economy*, 113, 6, 1217-1238.

Ellison G., Ellison S. (2009), "Search, Obfuscation and Price Elasticities on the Internet," *Econometrica*, 77, 2, 427-452.

Ellison G., Wolitzky A. (2009), "A Search Cost Model of Obfuscation," working paper.

Fudenberg D., Tirole J. (1991), "Perfect Bayesian Equilibrium and Sequential Equilibrium," *Journal of Economic Theory*, 53, 236-260.

Grossman S. (1981),"The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics*, 24(3),461-83.

Grossman S., Hart O. (1980),"Disclosure Laws and Takeover Bids," *Journal of Finance*, 35(2), 323-34.

Kamenica E., Gentzkow M. (2010), "Bayesian Persuasion," forthcoming in *American Economic Review*.

Kartik N. (2009), "Strategic Communication with Lying Costs," *Review of Economic Studies*, 76, 4, 1359-1395.

Mailath G.J., Okuno-Fujiwara M., Postlewaite A. (1993), "Belief-Based Refinements in Signaling Games," *Journal of Economic Theory*, 60, 2, 241-276.

Milgrom P. (1981), "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12, 380-91.

Milgrom P., Segal I. (2002), "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70, 2, 583-601.

Milgrom P., Shannon C. (1994), "Monotone Comparative Statics," *Econometrica*, 62, 1, 157-180.

Myerson R. (1983), "Mechanism Design by an Informed Principal," *Econometrica*, 51, 6, 1767-1797.

Perez-Richet E. (2012), "Competing with Equivocal Information," Working Paper.

Skreta V., Veldkamp L. (2009), "Ratings Shopping and Asset Complexity: A Theory of Ratings Inflation," *Journal of Monetary Economics*, 56, 5, 678-695.

Sobel J. (2007), "Signaling Games," in M. Sotomayor (ed.), *Encyclopedia of Complexity and System Science*, Springer, forthcoming.

Sokal A., Bricmont J. (1998), “Fashionable Nonsense: Postmodern Intellectuals’ Abuse of Science,” Picador USA, New York.

Umbhauer G. (1994), “Forward Induction, Consistency, Preplay Communication and ε-perturbations,” mimeo BETA. Université Louis Pasteur, Strasbourg, France.